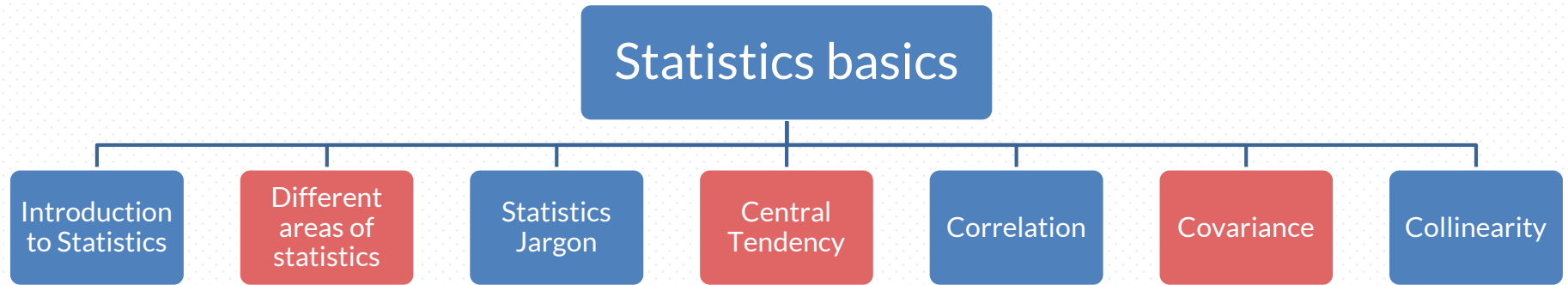




Basic Statistics

- MUQUAYYAR AHMED
DATA SCIENTIST

Agenda We Learnt...



Today's Agenda

```
graph LR; A[Today's Agenda] --- B[Hypothesis testing]; A --- C[ANOVA]; A --- D[Chi-square test];
```

Hypothesis testing

ANOVA

Chi-square test

Hypothesis Testing

- A statistical hypothesis is an assumption about a population parameter.
- Used to determine whether there is enough evidence in a sample of data to infer that a certain condition is true for the entire population.
- There are two types of statistical hypotheses.
 - Null hypothesis: denoted by H_0 , is usually the hypothesis that sample observations result purely from chance.
 - Alternative hypothesis: denoted by H_1 , is the hypothesis that sample observations are influenced by some non-random cause.
- p-value!!!

ANOVA

- ANOVA is a statistical technique used to compare the means of two or more groups of observations.
- ANOVA compares the means between the groups you are interested in and determines whether any of those means are significantly different from each other.
- Specifically, it tests the null hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

where μ = group mean and k = number of groups.

Chi-square Test

- Chi-square test is used to check the independencies between two variables.
- The test is applied when you have two categorical variables from a single population.
- Hope this examples helps
 - Is your manager's mood associated with the weather?

Mood \ Weather	Happy☺	SAD☹
Sunny	72%	28%
Rainy	72%	28%

- This appears to be no association between your manager's mood and the weather here because the row percentages are in same in each column.

Contd..

- Is your manager's mood associated with the weather?

Mood \ Weather	Happy☺	SAD☹
Sunny	82%	18%
Rainy	60%	40%

- There appears to be an association here because the row percentage is different in each column.